



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry

Schulz, Daniel ; Zanutelli, Vito Riccardo Tomaso ; Fischer, Jana Raja ; Schapiro, Denis ; Engler, Stefanie ; Lun, Xiao-Kang ; Jackson, Hartland Warren ; Bodenmiller, Bernd

Abstract: To build comprehensive models of cellular states and interactions in normal and diseased tissue, genetic and proteomic information must be extracted with single-cell and spatial resolution. Here, we extended imaging mass cytometry to enable multiplexed detection of mRNA and proteins in tissues. Three mRNA target species were detected by RNAscope-based metal in situ hybridization with simultaneous antibody detection of 16 proteins. Analysis of 70 breast cancer samples showed that HER2 and CK19 mRNA and protein levels are moderately correlated on the single-cell level, but that only HER2, and not CK19, has strong mRNA-to-protein correlation on the cell population level. The chemoattractant CXCL10 was expressed in stromal cell clusters, and the frequency of CXCL10-expressing cells correlated with T cell presence. Our flexible and expandable method will allow an increase in the information content retrieved from patient samples for biomedical purposes, enable detailed studies of tumor biology, and serve as a tool to bridge comprehensive genomic and proteomic tissue analysis.

DOI: <https://doi.org/10.1016/j.cels.2017.12.001>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-145495>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

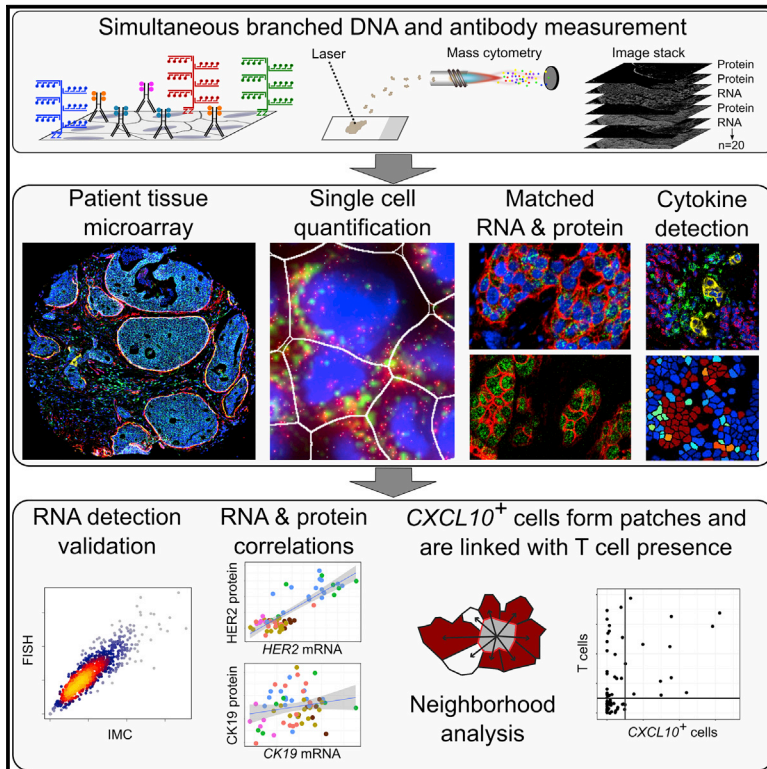
Originally published at:

Schulz, Daniel; Zanutelli, Vito Riccardo Tomaso; Fischer, Jana Raja; Schapiro, Denis; Engler, Stefanie; Lun, Xiao-Kang; Jackson, Hartland Warren; Bodenmiller, Bernd (2018). Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry. *Cell Systems*, 6(1):25-36.e5.

DOI: <https://doi.org/10.1016/j.cels.2017.12.001>

Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry

Graphical Abstract



Authors

Daniel Schulz,
Vito Riccardo Tomaso Zanotelli,
Jana Raja Fischer, ..., Xiao-Kang Lun,
Hartland Warren Jackson,
Bernd Bodenmiller

Correspondence

bernd.bodenmiller@imls.uzh.ch

In Brief

Here, we extend imaging mass cytometry to enable multiplexed detection of mRNA and protein in single cells in tissue sections. We show rigorous validation of the method and apply it to 70 samples from breast cancer patients. We investigate single-cell and population-based RNA-to-protein correlations in tissue and identify rare chemokine-expressing cells in the stroma that are associated with T cell abundance.

Highlights

- Imaging mass cytometry enables multiplexed RNA and protein detection *in situ*
- mRNA measurement in IMC enables detection of as little as 6–14 mRNA copies per cell
- Among patients, mRNA-to-protein ratios vary for *CK19* but not for *HER2*
- *CXCL10*-expressing cells form patches and are associated with T cell abundance

Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry

Daniel Schulz,¹ Vito Riccardo Tomaso Zanotelli,^{1,2} Jana Raja Fischer,¹ Denis Schapiro,¹ Stefanie Engler,¹ Xiao-Kang Lun,¹ Hartland Warren Jackson,¹ and Bernd Bodenmiller^{1,3,*}

¹Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland

²Systems Biology PhD Program, Life Science Zurich Graduate School, ETH Zurich and University of Zurich, Zurich, Switzerland

³Lead Contact

*Correspondence: bernd.bodenmiller@imls.uzh.ch

<https://doi.org/10.1016/j.cels.2017.12.001>

SUMMARY

To build comprehensive models of cellular states and interactions in normal and diseased tissue, genetic and proteomic information must be extracted with single-cell and spatial resolution. Here, we extended imaging mass cytometry to enable multiplexed detection of mRNA and proteins in tissues. Three mRNA target species were detected by RNAscope-based metal *in situ* hybridization with simultaneous antibody detection of 16 proteins. Analysis of 70 breast cancer samples showed that *HER2* and *CK19* mRNA and protein levels are moderately correlated on the single-cell level, but that only *HER2*, and not *CK19*, has strong mRNA-to-protein correlation on the cell population level. The chemoattractant *CXCL10* was expressed in stromal cell clusters, and the frequency of *CXCL10*-expressing cells correlated with T cell presence. Our flexible and expandable method will allow an increase in the information content retrieved from patient samples for biomedical purposes, enable detailed studies of tumor biology, and serve as a tool to bridge comprehensive genomic and proteomic tissue analysis.

INTRODUCTION

Characterization of the organization of cellular phenotypes, functions, and interactions in the context of tissues is key to our understanding of health and disease. The function of a tissue is defined by the cell types it contains, their arrangement (i.e., tissue morphology), and the state of each individual cell. The state of a cell, in turn, is defined by multiple networks that interact with each other to continuously adjust cell state according to internal and external inputs. Three network types that are interwoven to achieve cellular homeostasis are transcriptional networks, protein networks, and signaling networks. Simultaneous measurement of these networks *in situ* would allow one to derive quantitative models that enable understanding of these networks in a spatial context and thus enable study of many aspects of tissue biology.

Until recently only a few transcripts, proteins, or other molecules could be imaged at one time in tissues, but now several approaches allow for spatially resolved 'omics-type measurements (Bodenmiller, 2016). Immunofluorescence-based multiplexed protein epitope detection technologies such as cyclic immunofluorescence rely on cycles of epitope staining followed by quenching and restaining to overcome spectral overlaps of fluorophores (Gerdes et al., 2013; Lin et al., 2015). Alternatively, epitope-based imaging methods that employ a mass spectrometer for readout, such as multiplexed ion beam imaging and imaging mass cytometry (IMC), rely on the simultaneous staining and subsequent detection of up to 7 and 32 metal-labeled antibodies in tissue samples, respectively (Angelo et al., 2014; Bodenmiller, 2016; Giesen et al., 2014; Schapiro et al., 2017). Despite the power of these approaches, one common limitation is that the antibodies used must be comprehensively validated and optimized.

Methods based on *in situ* mRNA sequencing and encoded fluorescent *in situ* hybridization (FISH) probes have also been developed for spatial transcriptomics using fluorescence-based methods (Ke et al., 2013; Lee et al., 2014). These methods allow for the simultaneous detection of hundreds of distinct mRNAs under routine settings and in some cases over 1,000 transcripts (Chen et al., 2015). Targeted RNA detection methods using padlock probes, *in situ* hybridization chain reaction, and z-probes coupled to branched DNA amplification (RNAscope) also enable robust detection of RNA in tissue (Choi et al., 2014; Larsson et al., 2010; Wang et al., 2012) and have high signal-to-noise ratios (Battich et al., 2013), and their multiplexing capabilities are, among other things, limited by spectral overlaps of the detection reagents (Gaspar and Ephrussi, 2015; Wang et al., 2012). Although methods for the global measurement of the components of transcriptional or protein networks with spatial resolution in tissues are rapidly developing, approaches that enable mRNA, protein, and protein modification measurements in a highly multiplexed manner have, to our knowledge, so far not been presented.

Such methods, however, are necessary to study how transcriptional, protein, and signaling networks relate to each other. Many studies have investigated such relations in the form of RNA and protein-level correlations at a global scale in bulk samples (Liu et al., 2016). Based on these studies, it appears that protein expression can be largely explained by transcript abundance

(Jingyi and Biggin, 2015; Liu et al., 2016), and gene-specific conversion factors have recently been shown to increase RNA-protein correlations to 0.93 (Edfors et al., 2016). In certain cancer types, such as colon and rectal cancer, large variations in the correlation of RNA and protein abundances were observed across genes and patient samples (Zhang et al., 2014). The same study also showed that gene copy-number aberrations, which are among the leading causes of tumorigenesis (Stratton et al., 2009), are well correlated with mRNA levels but not always with protein levels, indicating the need for further investigations. In single cells, proof-of-principle approaches based on proximity ligation assays and DNA-tagged antibody sequencing indicate that RNA-to-protein correlations are typically poor, but such measurements are challenging and are restricted to few cells in suspension (Albayrak et al., 2016; Darmanis et al., 2016; Frei et al., 2016; Stoeckius et al., 2017). The relationship of RNA-to-protein levels *in situ* on the single-cell level and across tumor samples with copy-number alterations has not been studied so far.

Here, we present an approach for the simultaneous detection of proteins, protein phosphorylations, and transcripts using IMC. The approach is a modification of the RNAscope-based *in situ* hybridization protocol (Wang et al., 2012) coupled with antibody staining. We rigorously validated the approach in sections of human HeLa cell pellets and showed excellent agreement with FISH measurements. Furthermore, we characterized the mRNA expression of *HER2*, *CK19*, and *CXCL10* (also known as *IP-10*) in combination with 16 antibodies in 70 samples from breast cancer patients. The *HER2* gene is frequently genetically amplified in breast cancer patients, and *HER2* mRNA expression levels have been shown to be highly correlated to genetic status and protein levels (Vassilakopoulou et al., 2014; Wang et al., 2013), making *HER2* an ideal gene for validation of the system. *CK19* is a type-1 cytokeratin often expressed in breast cancer that can be used as a marker to detect disseminated tumor cells in lymph nodes by mRNA or protein detection methods, but the mRNA-to-protein correlation in tissues is unknown (Visser et al., 2008). *CXCL10* is a secreted protein that recruits T cells to the tumor microenvironment, yet its role in cancer needs to be further refined (Liu et al., 2011; Luster and Ravetch, 1987; Mulligan et al., 2013). We found that *HER2* mRNA and protein expression correlated well across patients on a population level, although this correlation was weak on the single-cell level in most cases. For *CK19*, we found a strong patient-dependent heterogeneity in mRNA-to-protein ratios. The expression of *CXCL10* was correlated with T cell presence in the tumor microenvironment, and, interestingly, cells expressing *CXCL10* were found to form clusters. We expect that the ability to simultaneously quantify mRNAs, proteins, and post-translational modifications will broaden our understanding of cellular networks in healthy and diseased tissues, enable measurement of markers not readily accessible via antibodies such as cytokines, and result in more accurate single-cell resolved tissue model generation to advance tissue biology and diagnostics.

RESULTS

RNA Detection with IMC

To enable RNA detection using IMC, we adapted the widely used RNAscope technique (Wang et al., 2012) for detection with metal

tags. The approach makes use of serial hybridization steps to build multiple, large DNA structures across an RNA molecule of interest (Figure 1A). Each of these structures can then be hybridized to hundreds of oligonucleotides bound to reporters. For mass-cytometric detection we conjugated oligonucleotides with metal-chelated polymer reporters and used these probes in the final hybridization step of the RNA detection protocol. After mRNA tagging by RNAscope-based *in situ* hybridization, an additional staining step with metal-labeled antibodies can be performed (Figure 1A, right) for the simultaneous detection of RNAs and proteins by IMC (Figure 1B).

Proof-of-principle experiments were carried out using sections of formalin-fixed and paraffin-embedded (FFPE) HeLa cell pellets. Sections were stained for mRNAs of three housekeeping genes *POLR2A*, *PPIB*, and *UBC* (Figure 1C). As a negative control we used probes designed to hybridize to the *DapB* transcript from *Bacillus subtilis*. Comparing the mean intensities per image for each measured transcript resulted in reproducible staining intensities with signal-to-noise ratios (S/N) of 18, 61, and 62 for *POLR2A*, *PPIB*, and *UBC*, respectively (Figure 1D; note that the S/N values of *PPIB* and *UBC* are similar due to slightly higher background signal for *UBC*). Comparing our measured expression levels with reported results from the literature (*POLR2A* expressed at low levels and *PPIB* and *UBC* expressed at high levels) revealed similar trends. When we compared for multiplex or single-plex detections of targets no obvious differences were observed in signals, indicating negligible cross-hybridization between RNA channels (Figure S1A).

Previously *POLR2A* transcripts were quantified using RNAscope and shown to be present at 14 copies per single HeLa cell with the section thickness used here (Anderson et al., 2016; Wang et al., 2012). For comparison of these numbers with our IMC data, we had to define mRNA signals from single cells. We used a DNA intercalator and a pan-keratin antibody to detect nuclei and cytoplasm, respectively, in combination with the three RNA species. Semi-supervised segmentation of the images was employed to obtain RNA data from individual cells (see STAR Methods). The IMC single-cell data showed that there were large variations in *POLR2A*, *PPIB*, and *UBC* amounts across cells (Figure 1E). *POLR2A* mRNA was detected with a mean signal of 207 ion counts per cell. To determine the detection limit of our approach we measured *TBP*, which is present at approximately six mRNA copies per cell. We were not able to detect *TBP* mRNA (Figure S1B), indicating that our detection threshold lies between 6 and 14 mRNA molecules per HeLa cell section. In summary, these results show that we can detect transcripts with IMC using an adapted RNAscope-based protocol with high sensitivity.

Validation Using FISH

The standard application of multiplexed RNAscope *in situ* hybridization is mRNA detection using fluorescence microscopy (Wang et al., 2012). Thus to benchmark the IMC-based RNA detection against the standard in the field, we devised an approach whereby mRNAs were simultaneously labeled with metals and fluorochromes for direct comparison of the labeling methods.

We applied a 1:1 mix of fluorochrome-labeled and metal-labeled oligonucleotides in the last hybridization step for the

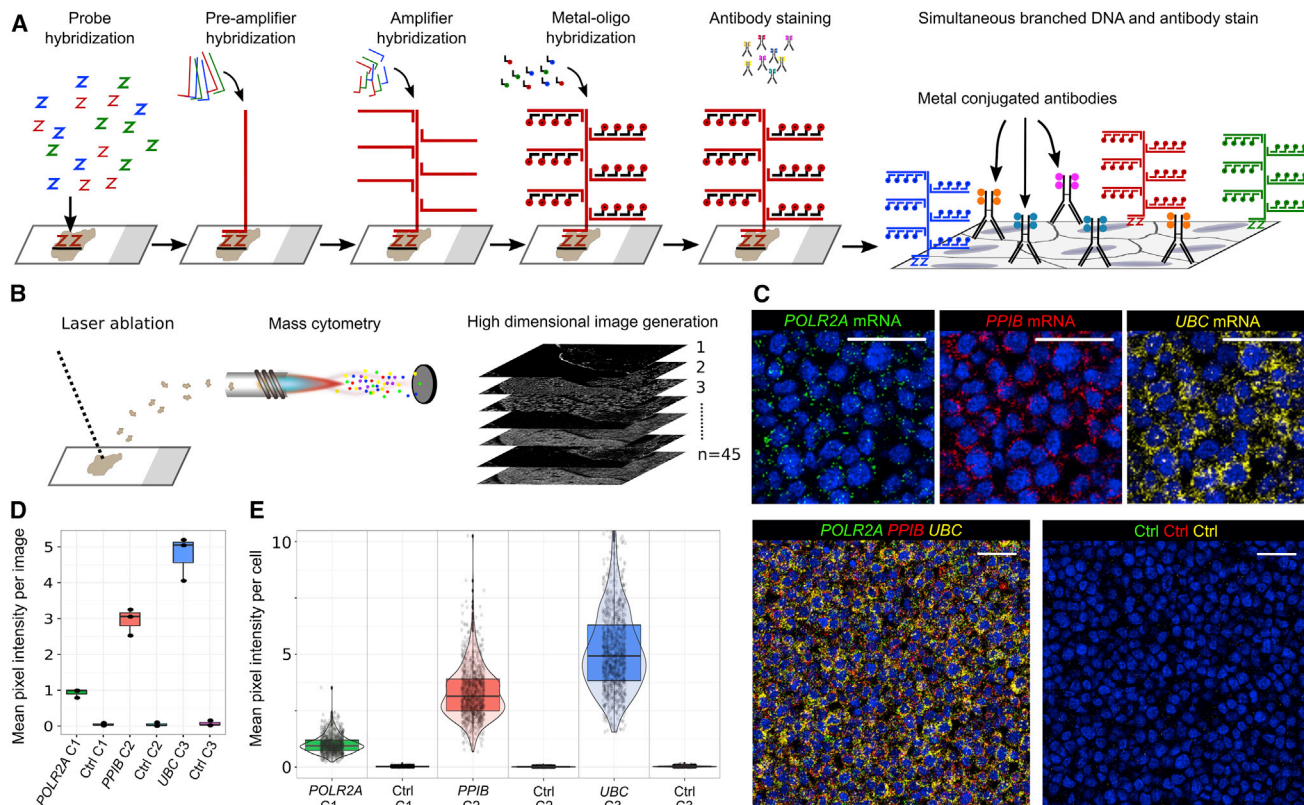


Figure 1. RNA Detection Using Imaging Mass Cytometry

(A) Up to 20 z-probe pairs per target RNA (up to three mRNAs may be detected simultaneously; red, green, blue) are incubated with the tissue sample. Through pre-amplifier and amplifier hybridization, a large DNA tree is assembled that is then hybridized by oligonucleotides labeled with metals. Metal-conjugated antibodies are added to enable co-detection of RNA and protein in single cells.

(B) After incubation with metal-labeled oligonucleotides for RNA detection and antibodies for protein detection, tissues are subjected to laser ablation and mass-cytometric measurement of the metal abundances. The metal abundances per laser shot (in pixels) are then assembled into a high-dimension image.

(C) Representative images stained for *POLR2A* in channel 1 (green, dynamic range 2–10 ion counts per pixel), *PPIB* in channel 2 (red, dynamic range 2–20 ion counts per pixel), and *UBC* in channel 3 (yellow, dynamic range 2–30 ion counts per pixel) are shown separately and overlaid. An image of the negative control probe set is shown on the bottom right for the three channels. The dynamic range for the negative control in each channel is 2–10 ion counts per pixel. Scale bars 50 μ m.

(D) Quantifications of the mean pixel intensities per image of three technical replicates for *POLR2A*, *PPIB*, and *UBC* and corresponding negative controls for each channel.

(E) Mean pixel intensities per cell after single-cell segmentation are plotted for *POLR2A*, *PPIB*, and *UBC* and corresponding negative controls. Box plots are overlaid with the single-cell data (black circles), and the single-cell data distributions are indicated.

detection of *PPIB* and *POLR2A* mRNA in HeLa cell sections (Figure 2A). Fluorescent images of the HeLa cell sections were recorded, and the same area was subsequently analyzed using IMC to detect metal abundances (Figure 2B). Overlay of data from the two imaging modalities enabled direct per-cell comparison of the ion counts from IMC with absolute copy numbers of RNA molecules from the fluorescent images (Figures 2C and 2D) (Battich et al., 2013). Comparison of the single-cell RNA copy numbers with the corresponding ion count signal per cell revealed a correlation of 0.89 for *PPIB* (Figure 2E). For the comparison of FISH and IMC detection, a pixel-level correlation was not feasible as FISH imaging was performed in wet samples and IMC imaging was performed after the samples were dried, which resulted in slight shifts of the imaged structures in IMC compared with FISH. For the *POLR2A* transcript, which is expressed at lower levels, we found a correlation of 0.80 between fluorescence and IMC (Figure 2F). IMC

and FISH signals were similarly distributed, and no signal accumulation occurred around zero. We concluded from these experiments that the adapted RNAscope-based *in situ* hybridization protocol yields results comparable with those for fluorescence microscopy, and thus is well suited for IMC-based single-cell RNA measurements.

The IMC RNA Detection Protocol Enables Simultaneous Multiplexed Detection of Proteins and Their Modifications

The many measurement channels available in IMC make it possible to interrogate the interplay of components of transcriptional, protein, and signaling networks in tissues in single cells by measuring mRNA and protein epitopes simultaneously. We therefore investigated simultaneous mRNA detection and antibody-based epitope detection with IMC in human FFPE breast cancer samples. The RNA detection protocol involves a

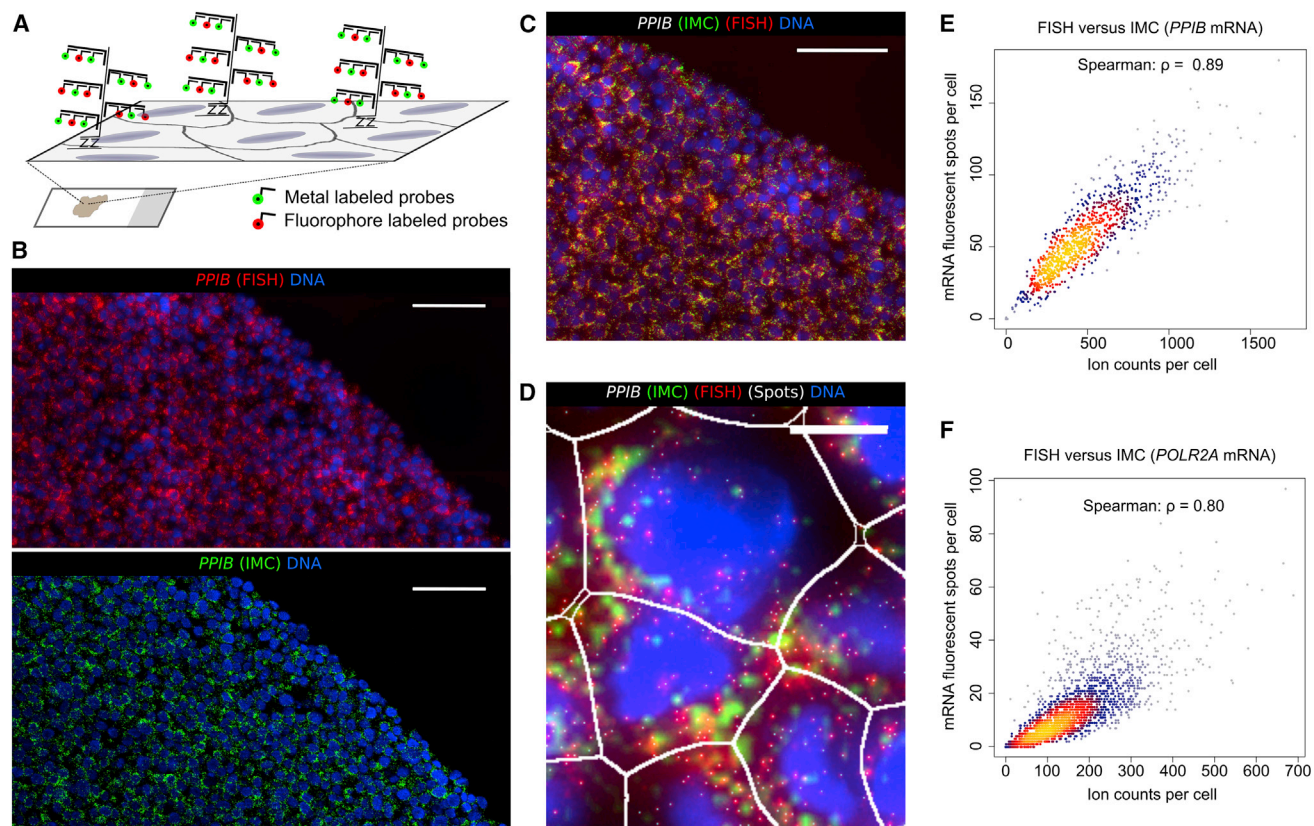


Figure 2. Validation of IMC-Based RNA Measurements Using FISH

(A) A 1:1 mix of metal-labeled and fluorophore-labeled oligonucleotide probes was applied to the sample to enable detection of the same transcript by fluorescence microscopy and IMC.

(B) False-color images from fluorescence microscopy (red, *PPIB*; blue, DAPI) and IMC (green, *PPIB*; blue, iridium DNA intercalator) of the same area are shown. Scale bars, 100 μm.

(C) Image of overlaid fluorescent and IMC images after registration (green, *PPIB* from IMC; red, *PPIB* from FISH). Scale bar, 100 μm.

(D) Zoom-in of overlay. White dots mark spots from the fluorescent image that were counted to obtain absolute RNA counts per cell (see [STAR Methods](#)). Cell outlines from segmentation are shown in white. Scale bar, 10 μm.

(E) Scatterplot of *PPIB* FISH mRNA spots versus IMC ion counts per cell. One representative replicate of the three technical replicates is shown.

(F) Scatterplot of *POLR2A* FISH mRNA spots versus IMC ion counts per cell. One representative replicate of two technical replicates is shown.

protease digestion step with multiple hybridization steps at 40°C that may destroy epitopes. We therefore compared antibody staining of 4-μm-thick serial sections of a breast cancer tissue that were processed using either a standard heat-induced epitope retrieval (HIER) protocol ([STAR Methods](#)) or the RNA detection protocol ([Wang et al., 2012](#)).

The images derived from the two protocols were similar, suggesting that protease treatment did not strongly interfere with detection of the antibodies used here ([Figure 3A](#)). We did, however, observe differences in CD44 staining. After HIER buffer antigen retrieval, CD44 staining was strongest in epithelial cells, whereas after the RNA protocol treatment the strongest staining was observed in stromal cells ([Figure S2A](#)). For a comprehensive and quantitative comparison we segmented the images to generate single-cell data and clustered the single-cell data based on the co-occurrence of marker expression ([Figure 3B](#)). Comparing the co-expression of cellular markers in images from the HIER protocol with co-expression in images from the RNA detection protocol revealed a correlation of 0.82, which increased to 0.92 when CD44 data were excluded. This

confirmed that the protease treatment used in our protocol does not destroy antibody-based epitope detection; however, the performance of each antibody to be used with the RNA detection protocol should be assessed.

Analysis of *HER2* and *CK19* mRNA and Protein Expression in Breast Cancer Tissues with Non-amplified and Amplified *HER2* Loci

Copy-number aberrations are among the main drivers of tumorigenesis ([Stratton et al., 2009](#)). How these copy-number aberrations affect transcript and the corresponding protein levels in single cells in tumors has so far not been analyzed ([Albayrak et al., 2016](#)). Here we quantified *HER2* and *CK19*, and *CXCL10* mRNA levels simultaneously with the analysis of 16 antibodies including antibodies against *HER2* and *CK19* protein ([Table S1](#)). These markers were measured in a previously described breast cancer tissue microarray (TMA) ([Carvajal-Hausdorf et al., 2015](#)). The amplification status of the *HER2* gene is known for each of the samples on the microarray ([Table S2](#)). In the 70 samples analyzed, *HER2* expression as

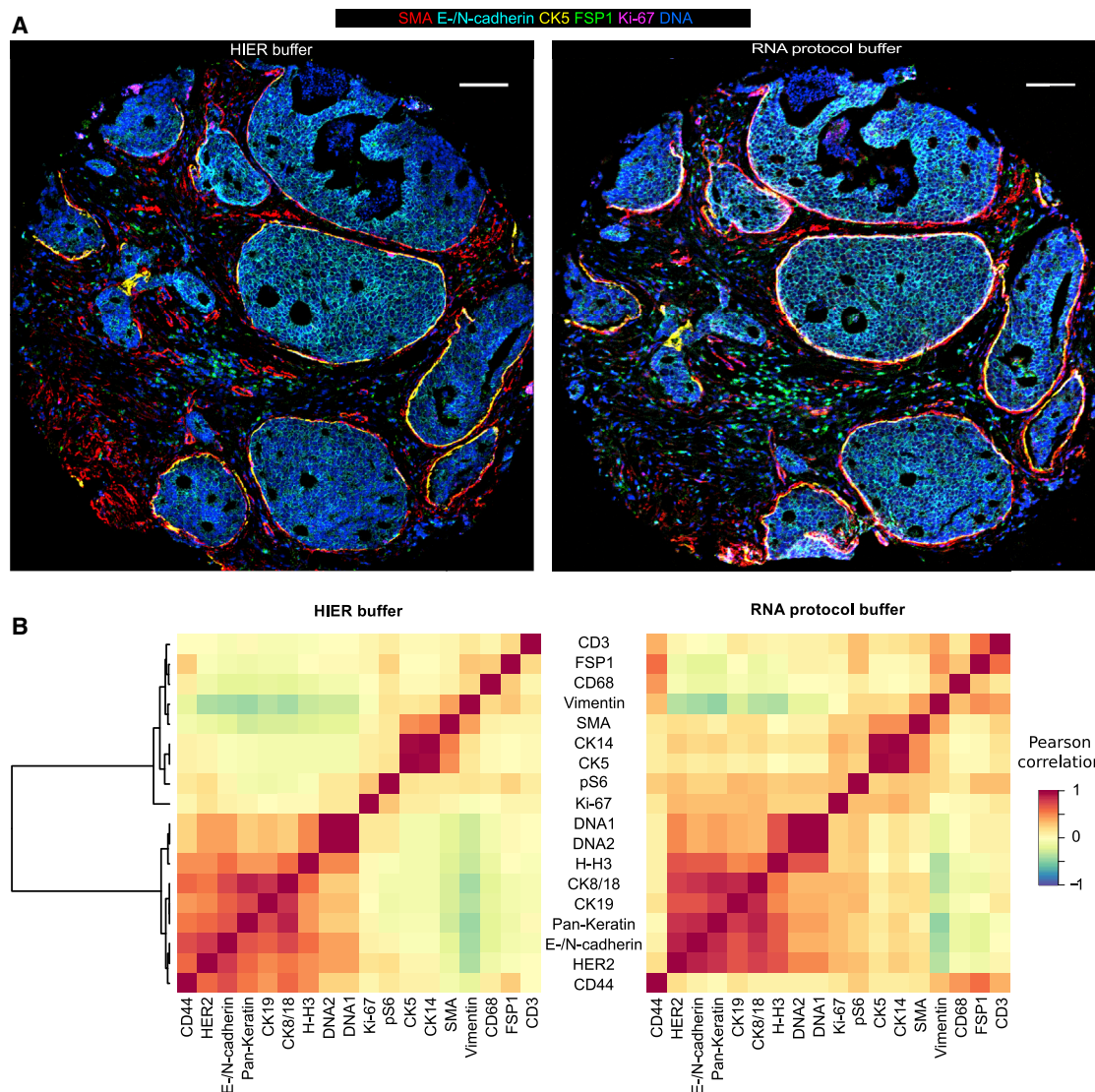


Figure 3. Multiplexed mRNA and Antibody Detection Using IMC

(A) False-color images for smooth muscle actin (red), E-/N-cadherin (cyan), Ki-67 (magenta), FSP1 (green), CK5 (yellow), and DNA (blue) antibody stains are shown for standard HIER and the RNA co-detection protocol. Consecutive tissue sections of the same region of the tumor were analyzed. Scale bars, 150 μ m. (B) Heatmap of Pearson correlations between signals in single cells for indicated markers in samples prepared with HIER buffer and RNA detection buffer. Correlations are shown for one biological replicate of a total of two biological and two technical replicates.

defined previously ranged from undetectable (score = 0) to strong *HER2* overexpression (score = 3+) (Carvajal-Hausdorf et al., 2015).

To verify the specificity of the mRNA detection probes used for human TMA analysis, we overexpressed and detected *CK19*, *CXCL10*, and *ERBB2* in A431 cells using the RNAscope-based IMC method (Figures S2B–S2D). We observed strong staining for *CK19*, *CXCL10*, and *HER2* only in cells in which the mRNAs were overexpressed, demonstrating specificity of the probes (Figure S2E).

After staining and data acquisition of the *HER2* TMA, visual inspection of the images revealed heterogeneous staining for *HER2* and *CK19* mRNAs and proteins across single cells (Figure 4A). Analysis of the *HER2* mRNA expression in epithelial cells

revealed that 21 of the 26 samples with *HER2* amplification showed significant upregulation of the *HER2* mRNA compared with control tissues (Figure 4B). Of the amplified samples, 16 had levels of *HER2* mRNA that were more than 10-fold higher than levels in control tissues. Of the 35 samples without *HER2* amplification, 21 also had significant upregulation of *HER2* mRNA compared with control tissues, albeit with a lower (2- to 7-fold) induction than in the majority of *HER2*-amplified samples (Figure 4B).

For *HER2* protein, 19 of the 26 samples with *HER2* amplification showed significant expression of *HER2* compared with control tissues. Three of these 19 *HER2*-amplified samples showed significant overexpression of the *HER2* protein but their levels of *HER2* mRNA were lower than in most other amplified samples

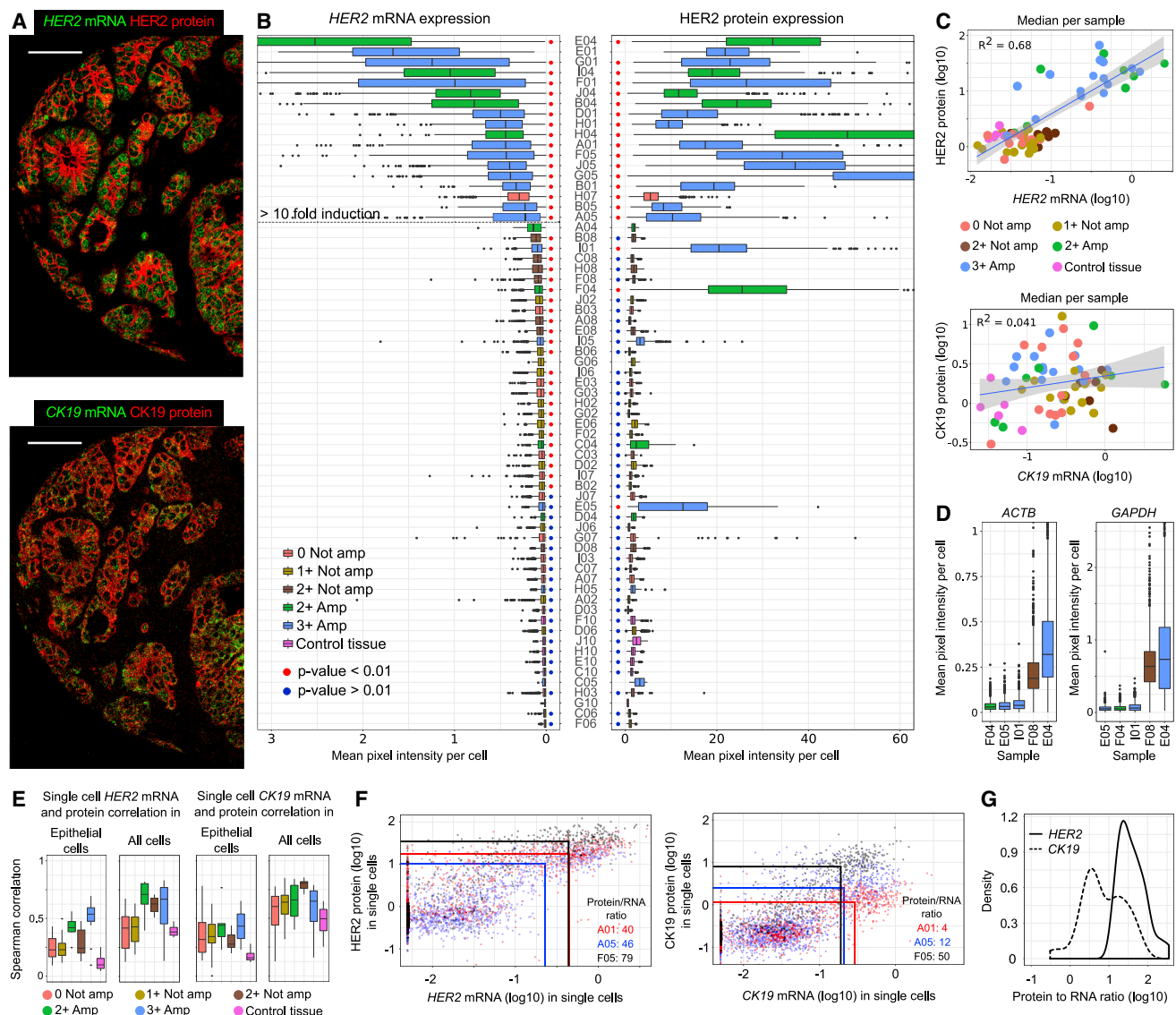


Figure 4. Population and Single-Cell *HER2* and *CK19* mRNA and Protein Correlation Analyses

(A) Single false-color image of a representative tissue sample (E04) displaying *HER2* mRNA (green) and protein (red) (top) and *CK19* mRNA (green) and protein (red) (bottom). Scale bars, 100 μ m.

(B) Box plots of *HER2* mRNA (left) and protein (right) expression in single epithelial cells from each sample. Samples were sorted by increasing expression of *HER2* mRNA. Red dots indicate that data were significantly different from control tissue data (see STAR Methods). Blue dots indicate that the data were not significantly different from the control tissues, and no dot means that the sample contained less than 50 epithelial cells.

(C) The median expression in all Pan-Keratin-positive epithelial cells for each sample is shown for *HER2* mRNA and protein (top) and *CK19* mRNA and protein (bottom).

(D) Expression of housekeeping genes *ACTB* and *GAPDH* in the three samples in which strong *HER2* protein but no *HER2* mRNA was detected, compared with one sample in which mRNA and proteins were detected and one sample in which no *HER2* mRNA and protein were detected.

(E) Spearman correlations of mRNA and protein levels for single epithelial cells for *HER2* and *CK19*. Samples were grouped by pathological scoring and genetic status of *HER2*.

(F) Single-cell scatterplots of samples A01, A05, and F05 for *HER2* (left) and *CK19* (right) mRNA and protein. The colored horizontal and vertical lines equal the respective median protein and median mRNA values of the epithelial population in each sample, which were used to calculate the protein-to-RNA ratios.

(G) Density distributions of the *HER2* (solid line) and *CK19* (dashed line) protein-to-RNA ratios from the data in (C).

Data in (C) and (D) are colored according to pathological scoring and genetic status of *HER2* locus: 0, red; 1+, not amplified, yellow; 2+, amplified, green; 2+, not amplified, brown; 3+, amplified, blue; non-diseased control tissue, pink.

(Figure 4B). Neither mRNA nor protein measurements fully reflected the 2+ and 3+ pathological scoring of the samples. For example, samples scored as *HER2* 2+ and not genetically ampli-

fied showed generally low levels of *HER2* mRNA and protein (Figure 4B). The discrepancy between the pathological scoring and our measurement is likely due to use of different antibodies

and evaluation of different areas of the tumor. It should be noted that the pathological HER2 scoring was also not reproduced in previous studies of this TMA (Carvajal-Hausdorf et al., 2015).

The correlation of the median signals of *HER2* mRNA and protein for all samples was 0.68 (Figure 4C). In the same samples, the same analysis for *CK19* revealed essentially no correlation (Spearman correlation 0.16) despite a dynamic range similar to that of *HER2* (Figure 4C). Across all samples for *HER2* we found an approximately 8-fold increase in protein for a 10-fold increase in RNA, whereas *CK19* protein levels varied by 10-fold for samples with virtually identical mRNA levels. Among our cohort of samples E05, F04, and I01 stood out, as they showed high *HER2* protein levels but no or very low *HER2* mRNA expression. Therefore, we also investigated the expression of *ACTB* and *GAPDH* housekeeping genes in serial sections of these samples and found no expression of *ACTB* or *GAPDH* in these three samples, suggesting that in these cores the mRNA was degraded (Figure 4D).

In sum, these results show that our approach enables the multiplexed analysis of transcript and protein levels in tissue samples. In a majority of cases (but not always) the mRNA and protein levels of *HER2* reflected the amplification status. The population-level analysis also revealed a good correlation between mRNA and protein levels for *HER2* but not for *CK19* in the same cells, suggesting that these two genes have different post-transcriptional regulatory mechanisms.

Single-Cell Correlations of *HER2* and *CK19* mRNAs with Protein Levels

Population-based analysis can mask relationships that exist on the single-cell level within a sample (Altschuler and Wu, 2010). Thus, we next examined the correlation between mRNA and protein abundances of *HER2* and *CK19* on the single-cell level. Both *HER2* and *CK19* are expressed in epithelial cells, so we first studied the single-cell correlation in the epithelial cell compartment.

For *HER2*, weak correlations between mRNA and protein levels were observed for the amplified 2+ and 3+ samples with correlation values of 0.43 and 0.53, respectively (Figure 4E). For *CK19* the correlation of mRNA and protein in epithelial cells was independent of the amplification status of *HER2*, with correlations ranging from 0.28 to 0.45 across all sample groups (Figure 4E). Analyses including the stromal cells showed higher correlations for *HER2* and *CK19* than when epithelial cells were considered alone, ranging from 0.42 to 0.67 for *HER2* and from 0.56 to 0.81 for *CK19* (Figures 4E, S3, and S4). Interestingly, we found that while the single-cell mRNA and protein correlations for *HER2* and *CK19* were somewhat similar, the ratios of mRNA to protein for *CK19* varied considerably among samples even though the same samples had very little variation in their *HER2* mRNA-to-protein ratio (Figure 4F). Analysis of the ratios across all samples revealed that the variation of mRNA-to-protein ratios was 1.8-fold higher for *CK19* than for *HER2* and showed a somewhat bimodal distribution (Figure 4G). This explains at least partially the discrepancy between the population-based mRNA and protein correlations between *HER2* and *CK19* (Figure 4C). From the combined observations of either *CK19* or *HER2* mRNA expression (Figures S3 and S4), we observed only 7 samples (including E05, F04, and I01) that had

very low levels of mRNA, suggesting that these cores may have suffered from mRNA degradation. This corresponded to 11% of all samples, suggesting that the majority of samples in this cohort were intact.

In summary, the single-cell correlations between mRNA and protein levels for *HER2* and *CK19* were generally weak in epithelial cells, indicating that we cannot predict protein levels from mRNA levels in these cells. However, mRNA presence, with some exceptions, always resulted in the corresponding protein presence as indicated by the analysis of all cell types. The single-cell data also suggest sample-specific differences in the extent of mRNA-to-protein ratio for *CK19*, whereas the ratio of mRNA to protein was more consistent across samples for *HER2*.

Cell Phenotype Analysis Identifies Rare Cells that Express *CXCL10* mRNA

CXCL10 is produced upon interferon- γ (IFN- γ) stimulation and is secreted by various cell types into the extracellular matrix, where it functions as a chemoattractant for T cells (Luster and Ravetch, 1987). To identify cell phenotypes that produce *CXCL10* and to gain a global overview of all cell phenotypes present in the 81,974 human breast cancer cells in our dataset, we used the PhenoGraph algorithm to assign cells to phenotype clusters (Levine et al., 2015).

PhenoGraph clustering identified four major phenotypic branches: (1) immune cells and other cells residing in the stroma, (2) basal epithelial cells and myoepithelial cells, (3) epithelial cells not expressing *HER2*, and (4) epithelial cells expressing *HER2* (Figure 5A). Within these branches were cells that overexpress *HER2* protein (clusters 14, 16, and 32), cells that overexpress *HER2* mRNA and protein (clusters 9 and 10), and two clusters with high levels of *HER2* mRNA but low *HER2* protein levels (clusters 27 and 31). PhenoGraph also revealed two clusters of T cells (as indicated by high levels of CD3), one of which co-expressed CD44 and FSP1 (potentially memory T cells [Weathery et al., 2015], cluster 4) and one solely expressing CD3 (cluster 33). Three CD68-expressing clusters (17, 18, and 22) were identified, which are monocytes, macrophages, and/or fibroblasts (Duncan et al., 2002; Österreicher et al., 2011; Pilling et al., 2009). For additional clusters of stromal cells no clear cell type assignment was possible given the markers used in our analysis.

Importantly, PhenoGraph identified a cluster of 2,300 cells characterized by the expression of *CXCL10* mRNA. The *CXCL10* cluster was also characterized by expression of different markers from the stromal compartment, indicating mixed sources of *CXCL10* expression (Figure 5A, cluster 20). Close inspection of the *CXCL10*-positive cluster 20 revealed that the cluster contained cells with medium and high levels of *CXCL10* expression (Figure S5A). Subsequent manual gating identified 706 cells with high *CXCL10* expression (0.86% of all cells) (Figure S5A). Clustering of these cells using PhenoGraph revealed that the majority of cells expressing high levels of *CXCL10* are macrophages, fibroblasts, T cells, and other unidentified cells in the stroma; only about 20% of cells with high *CXCL10* expression were epithelial cells (Figure S5B). Thus, the mRNA and protein IMC data could be used to cluster cells into known subsets in the tumor microenvironment including T cells, macrophages, and different lineages of breast cancer

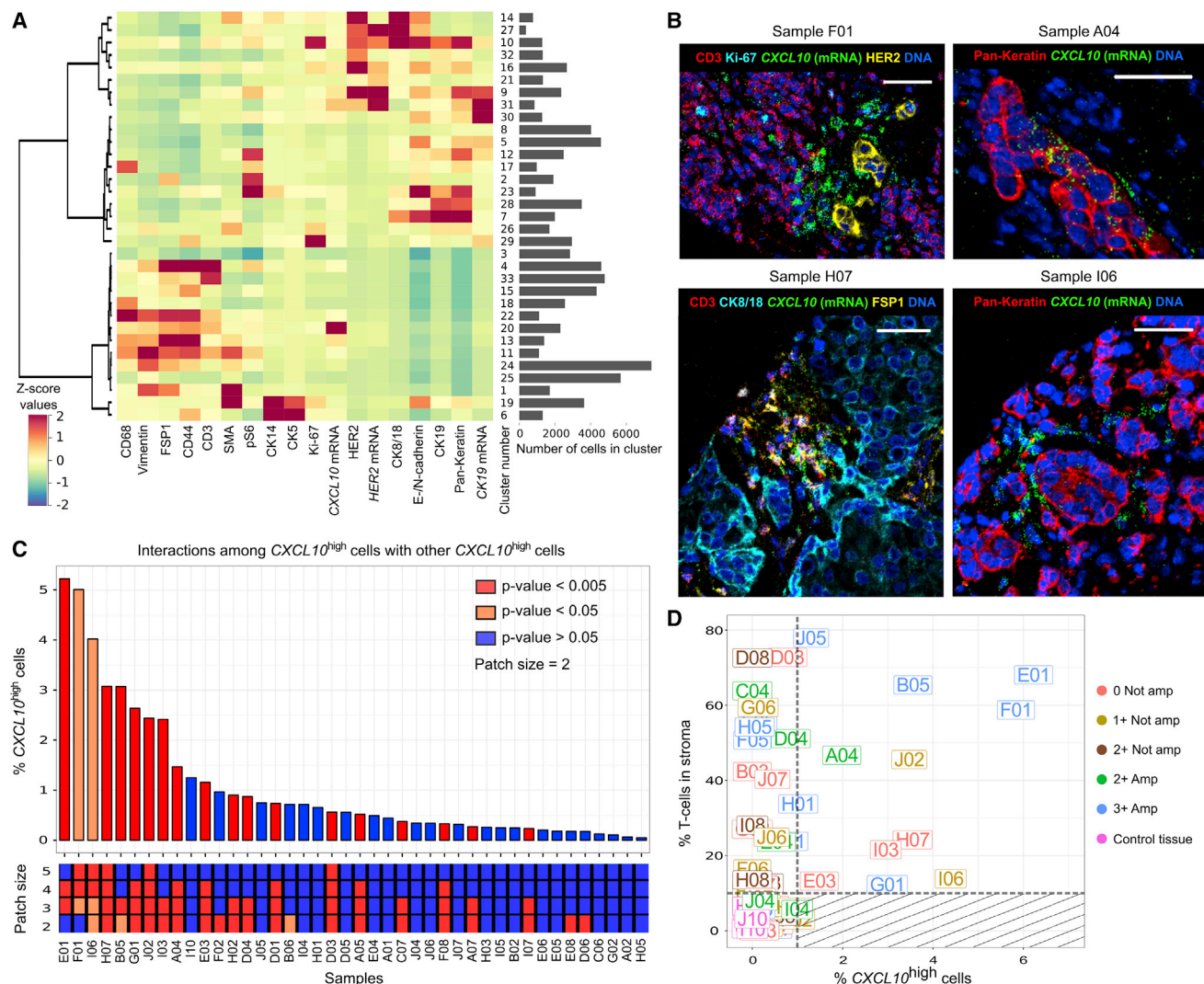


Figure 5. mRNA Analysis Coupled with Cell Neighborhood Analysis Reveals Correlation between CXCL10 Expression and T Cell Frequency

(A) Heatmap of mean marker expression for PhenoGraph clusters of all samples. Cell numbers in each cluster are plotted on the right side of the heatmap. (B) Example false-color images of four samples that contain cells of cluster-20 CXCL10-expressing cells. Scale bars, 50 μ m. (C) For samples with CXCL10-expressing cells, the percentages of CXCL10^{high} cells are plotted. Each bar is colored according to neighborhood analysis indicating whether CXCL10^{high} cells are clustered in patches of at least three cells (red, $p < 0.005$; light red, $p < 0.05$; blue, not significant [$p \geq 0.05$]). Below are the neighborhood analysis results presented for standard histoCAT neighborhood interactions among CXCL10^{high} cells (cluster size = 2) and for larger patches consisting of minimally 3, 4, or 5 cells. (D) Percent total CXCL10^{high} cells versus percent stromal T cells for every sample. Data are colored according to pathological scoring of HER2 protein and genetic status of HER2 locus: 0, red; 1+, not amplified, yellow; 2+, amplified, green; 2+, not amplified, turquoise; 3+, amplified, blue; non-diseased control tissue, pink. Gray shaded area indicates the region devoid of CXCL10^{high} cells (<1%) and T cells (<10%).

cells, and suggest that cells residing in the stroma are the major, but not exclusive, producers of CXCL10.

CXCL10-Expressing Cells Form Patches and Correlate with an Increased T Cell Presence

CXCL10 and its receptor CXCR3 have been comprehensively studied, and expression of these two factors has mostly been associated with poor survival and metastasis in different cancer types, although anti-tumor effects through immunomodulation have also been described (Bai et al., 2016; Bronger et al., 2016; Liu et al., 2011; Mulligan et al., 2013; Wightman et al.,

2015). Visual inspection of images with many cells expressing high levels of CXCL10 confirmed that CXCL10^{high} cells reside mostly in the stroma and seemed unevenly distributed (Figure 5B). To explore the spatial distribution of cells that express high levels of CXCL10 and their neighborhoods, we used our recently developed software histoCAT (Schapiro et al., 2017). A permutation test implemented in histoCAT determines whether cell types significantly more or less often neighbor each other than expected by chance. This test revealed no significant interactions of CXCL10^{high} cells with other cell types, but CXCL10^{high} cell interactions with other CXCL10^{high} cells were significant,

which can be indicative of a patchy distribution (Figure 5C, bottom; patch size = 2). We thus extended the permutation test to determine whether particular cell types form patches. We defined patches as motifs in which a cell of a certain phenotype neighbors at least two other cells of the same phenotype. We then used the permutation test to identify images with significantly increased frequencies of this motif, indicating an occurrence of the tested cell type in patches. Applying this test to the *CXCL10*^{high} cells in all images revealed a patchy distribution of *CXCL10*^{high} cells across many images, especially in those with high numbers of *CXCL10*^{high} cells (Figure 5C). Gradually increasing the minimum number of cells in the patch motif to five cells revealed those images with bigger clusters of *CXCL10*^{high} cells (Figure 5C, bottom; cluster size 3–5). This confirmed that *CXCL10*^{high} cells tend to cluster together and may form locally distinct microenvironments.

Since *CXCL10* is a chemoattractant for T cells (Khan et al., 2000), we next analyzed whether there was a relationship between the presence of *CXCL10*^{high} cells and the presence of T cells in the same sample. We compared the frequency of *CXCL10*^{high} cells with the frequency of T cells (CD3-positive gate, Figure S5C) in the stroma of all samples (Figure 5D). This analysis revealed samples with stromal T cell frequencies of up to 80%, and for samples with high numbers of *CXCL10*^{high} cells we always observed an enrichment of T cells in the stroma. Importantly, although images with high T cell frequencies but no *CXCL10*^{high} cells were present, we did not identify samples with *CXCL10*^{high} cells and no T cells (Figure 5D, gray shaded box). The absence of T cells when there are no *CXCL10*^{high} cells (Figure 5D, gray shaded box) was also found to be significant ($p = 0.001$, Barnard test), indicating the functionality of *CXCL10* expression and attraction of T cells by *CXCL10*^{high} cells. These data also display how the mRNA-based measurements can be used to detect markers that are difficult to detect with antibodies in tissues and how, when coupled with automated detection of cellular patches, the data can be used to define local microenvironments driven by a factor of interest.

DISCUSSION

Here we describe the validation and application of a strategy for multiplexed simultaneous RNA and protein measurements in single cells in human FFPE tissue by IMC. Co-detection of transcripts and proteins in tissue will be a powerful tool for studies of mRNA, protein, and signaling network relationships across a wide range of different cell types in a spatial context in tissue, with broad applicability to biological and biomedical questions. Additionally, mRNA readouts may be used to detect cells that produce secreted factors and targets for which an appropriate antibody is not available. Our comparison of IMC with FISH revealed very good agreement on a cellular level and a detection limit between 6 and 14 RNA copies in cross-sections of cells.

Our method is a modification of a widely used approach for mRNA detection using RNAscope (Wang et al., 2012) that enables detection of metal tags by mass cytometry. In this work, we detected 3 mRNAs and 16 proteins simultaneously with single-cell resolution in FFPE tissue sections. In our overexpression experiments (Figure S2) we observed single cells with staining

intensities 40-fold higher than what we observed here for *UBC*, which is endogenously expressed at high levels (Figure 1E). Such high expression levels suggest that many more probes can be accommodated and indicate that higher multiplexing will not be limited by the cellular space.

The RNA co-detection protocol presented here involves a protease digestion step that may interfere with epitope detection for certain antibodies. We did observe different staining patterns for CD44 after standard HIER buffer and after the RNA protocol treatment. Differences may be due to alterations in the accessibility of the many different isoforms of CD44 after the protease digestion step (Mackay et al., 1994). Like all antibody-based approaches, our method is limited by antibody availability and specificity. mRNA signals are low compared with protein signals, and careful panel design is required to achieve optimal detection and avoid spillover from antibody channels (Takahashi et al., 2016).

Here we investigated the relationship between two transcripts (*CK19* and *HER2*) and the corresponding proteins in 70 breast tumor samples. On a population level, we found good correlations for *HER2* mRNA and protein but poor correlation between *CK19* mRNA and protein. Some amplified samples, however, strongly expressed *HER2* protein but only weakly expressed *HER2* mRNA. In three samples that had prominent expression of *HER2* protein but not *HER2* mRNA (E05, F04, I01), RNA appeared to be degraded. Future studies should incorporate either up-front or simultaneous testing for RNA integrity through the detection of housekeeping gene expression to avoid false negatives due to RNA degradation. Analysis of the combined expression of *CK19* and *HER2* mRNA in this study suggest that only a minority of samples (11%) have suffered RNA degradation. The difference between *HER2* and *CK19* mRNA and protein correlations is therefore likely due to different mechanisms of regulation of gene expression that could be rooted in transcription, translation, or protein and mRNA stabilities. Alternatively, the differences across patient samples we identified in mRNA-to-protein level ratios could be due to patient-specific deregulation that might be reflected in clinical features. Such comparisons have never been done and will be closely inspected in our future studies. Importantly, the relationship between mRNA and protein levels for *HER2* and the lack thereof for *CK19* was found within the same cells, thereby strengthening our conclusions and demonstrating the power of multiplexed single-cell measurements.

We also investigated the expression of the chemokine *CXCL10* in tumor tissue sections from 70 breast cancer patients. *CXCL10* has been reported to have pro-tumorigenic and tumor-suppressive functions (Bai et al., 2016; Bronger et al., 2016; Duruisseaux et al., 2017; Liu et al., 2011; Peng et al., 2015; Wightman et al., 2015). Interestingly, in breast cancer samples from patients with known genetic profiles and clinical outcomes (Györfy et al., 2010), expression of the *CXCL10* transcript on a population level was associated with slightly but significantly poorer prognosis in estrogen receptor-positive tumors (Figure S5D) (Györfy et al., 2010). In tumors negative for both estrogen and progesterone receptors, expression of *CXCL10* was significantly associated with longer relapse-free survival (Figure S5E) (Györfy et al., 2010). These data suggest that *CXCL10* mRNA has potential as a biomarker, and further studies

are needed to elucidate its function in different tumor settings. *CXCL10* mRNA was not expressed in all breast cancer patient samples examined and was expressed only in a small percentage of cells. We observed most of the mRNA expression in stromal cells, in contrast to previous studies in which *CXCL10* protein staining was often observed in the epithelial compartment (Bai et al., 2016; Duruisseaux et al., 2017; Mulligan et al., 2013; Peng et al., 2015).

Interestingly, *CXCL10*-positive cells were often located in patches, suggesting that these cells attract each other, induce expression in neighbors, or are found in a specific microenvironment. That transcripts such as *CXCL10* may be used to infer properties of the environment could be validated through measurements of additional factors such as IFN- γ , the key cytokine that leads to *CXCL10* expression (Luster and Ravetch, 1987). Information on active transcription pathways suggestive of inflammatory events could be used to refine phenotypes of cells in the tumor microenvironments. *CXCL10* is a chemoattractant for T cells (Khan et al., 2000), and we observed a correlation between T cell frequencies in stroma and frequencies of cells expressing high levels of *CXCL10* mRNA. The epitopes detected in our study were not sufficient to reveal whether increased levels of *CXCL10* correlated with increased levels of cytotoxic T cells, CD4⁺ type 1 T helper (Th₁) cells, or Th₂ cells, all of which are known to be attracted through *CXCL10* (Groom and Luster, 2011). This will be addressed in future studies along with the question of which factors precisely control T cell phenotypes after recruitment into the tumor microenvironment.

In summary, the IMC protocol for simultaneous detection of RNA and protein targets described here enables comprehensive and detailed cellular, phenotypic, and functional characterizations of single cells in FFPE tissues. The approach is highly robust and flexible, shown also by another group that in parallel to this study applied the RNAscope-based RNA detection to cell lines using mass cytometry (Mavropoulos et al., 2017). The RNA detection protocol is scalable, and higher multiplexing will soon be possible. Our approach will guide the development of tissue architectural maps of health and disease in unprecedented detail and comprehensiveness. The ability to link transcript, protein, and signaling networks in tissues with spatial resolution will improve our understanding of tissue biology and will enable assessment of drug efficacy, discovery of novel routes for intervention, and identification of informative biomarkers, and thus has the potential to become a pillar for precision medicine approaches.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Biological Material
- METHOD DETAILS
 - Experimental Design
 - Antibody Conjugation
 - Oligonucleotide Conjugation

- Sample Pre-treatment
- IMC and FISH Comparison
- RNA and Protein Staining
- Overexpression of *CK19* and *CXCL10* and *ERBB2*
- Data Acquisition

● QUANTIFICATION AND STATISTICAL ANALYSIS

- Image Generation
- Pixel Classification in Ilastik for Segmentation
- Single Cell Segmentation
- Data Analysis
- Statistical Tests and Data Exclusion
- Neighborhood Analysis Using histoCAT
- RNA FISH and IMC Data Analysis

● DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and two tables and can be found with this article online at <https://doi.org/10.1016/j.cels.2017.12.001>.

ACKNOWLEDGMENTS

We thank Emily Park, Yuling Luo, and Xiao-Jun Ma from Advanced Cell Diagnostics for providing reagents for RNA detection and fruitful discussion. We further thank David Rimm for providing the HER2 TMA for this study and critically reading the manuscript, as well as Zuzana Varga, our collaborator at the University Hospital Zurich, for providing us with samples for protocol testing and comparison. We also thank Raul Catena for cutting tissue sections. B.B.'s research is funded by an SNSF R'Equip grant, an SNSF Assistant Professorship grant, the SystemsX Transfer Project "Friends and Foes", the SystemsX MetastasiX and PhosphoNetX grant, NIH grant (UC4 DK108132), and the European Research Council (ERC) under the European Union's Seventh Framework Program (FP/2007-2013)/ERC grant agreement no. 336921. D. Schulz was funded by an EMBO fellowship (ALTF 970-2014) co-funded by the European Commission (LTFCOFUND2013, GA-2013-609409) and the Forschungskredit of the University of Zurich, grant no. FK-17-115. D. Schapiro was supported by the Forschungskredit of the University of Zurich, grant FK-74419-01-01, and the BioEntrepreneur-Fellowship of the University of Zurich, reference no. BIOEF-17-001.

AUTHOR CONTRIBUTIONS

D. Schulz and B.B. designed the study. D. Schulz performed all experiments with help from S.E., X.-K.L., and H.W.J. D. Schulz. and V.R.T.Z. performed data analysis. D. Schulz, J.F., V.T.R.Z., and D. Schapiro adapted neighborhood analysis. D. Schapiro wrote code to overlay IF and IMC images and D. Schapiro and J.F. implemented spot detection in histoCAT. D. Schulz and B.B. wrote the manuscript. All authors commented on and edited the final version of the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 15, 2017

Revised: November 6, 2017

Accepted: December 1, 2017

Published: December 27, 2017

REFERENCES

- Albayrak, C., Jordi, C.A., Zechner, C., Lin, J., Bichsel, C.A., Khammash, M., and Tay, S. (2016). Digital quantification of proteins and mRNA in single mammalian cells. *Mol. Cell* 61, 914–924.
- Altschuler, S.J., and Wu, L.F. (2010). Cellular heterogeneity: do differences make a difference? *Cell* 141, 559–563.

- Anderson, C.M., Zhang, B., Miller, M., Butko, E., Wu, X., Laver, T., Kernag, C., Kim, J., Luo, Y., Lamparski, H., et al. (2016). Fully automated RNAscope in situ hybridization assays for formalin-fixed paraffin-embedded cells and tissues. *J. Cell. Biochem.* **117**, 2201–2208.
- Angelo, M., Bendall, S.C., Finck, R., Hale, M.B., Hitzman, C., Borowsky, A.D., Levenson, R.M., Lowe, J.B., Liu, S.D., Zhao, S., et al. (2014). Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442.
- Bai, M., Chen, X., and Ba, Y.I. (2016). CXCL10/CXCR3 overexpression as a biomarker of poor prognosis in patients with stage II colorectal cancer. *Mol. Clin. Oncol.* **4**, 23–30.
- Battich, N., Stoeger, T., and Pelkmans, L. (2013). Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat. Methods* **10**, 1127–1133.
- Bodenmiller, B. (2016). Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell Syst.* **2**, 225–238.
- Bronger, H., Singer, J., Windmüller, C., Reuning, U., Zech, D., Delbridge, C., Dorn, J., Kiechle, M., Schmalfeldt, B., Schmitt, M., and Avril, S. (2016). CXCL9 and CXCL10 predict survival and are regulated by cyclooxygenase inhibition in advanced serous ovarian cancer. *Br. J. Cancer* **115**, 553–563.
- Carvajal-Hausdorf, D.E., Schalper, K.A., Pusztai, L., Psyrri, A., Kalogeras, K.T., Koutala, V., Fountzilas, G., and Rimm, D.L. (2015). Measurement of domain-specific HER2 (ERBB2) expression may classify benefit from Trastuzumab in breast cancer. *J. Natl. Cancer Inst.* **107**, <https://doi.org/10.1093/jnci/djv136>.
- Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090.
- Choi, H.M.T., Beck, V.A., and Pierce, N.A. (2014). Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano* **8**, 4284–4294.
- Couzens, A.L., Knight, J.D.R., Kean, M.J., Teo, G., Weiss, A., Dunham, W.H., Lin, Z.-Y., Bagshaw, R.D., Sicheri, F., Pawson, T., et al. (2013). Protein interaction network of the mammalian hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci. Signal.* **6**, rs15.
- Darmanis, S., Gallant, C.J., Marinescu, V.D., Niklasson, M., Segerman, A., Flamourakis, G., Fredriksson, S., Assarsson, E., Lundberg, M., Nelander, S., et al. (2016). Simultaneous multiplexed measurement of RNA and proteins in single cells. *Cell Rep.* **14**, 380–389.
- Duncan, M.C., Costaguta, G., and Payne, G.S. (2002). Vimentin is secreted by activated macrophages. *Nat. Cell Biol.* **5**, 77–81.
- Duruisseau, M., Rabbe, N., Antoine, M., Vieira, T., Poulot, V., Cadranet, J., and Wislez, M. (2017). Pro-tumoral CXCL10/CXCR3-A autocrine loop in invasive mucinous lung adenocarcinoma. *ERJ Open Res.* **3**, <https://doi.org/10.1183/23120541.00047-2016>.
- Edfors, F., Danielsson, F., Hallström, B.M., Käll, L., Lundberg, E., Pontén, F., Forsström, B., and Uhlén, M. (2016). Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol. Syst. Biol.* **12**, 883.
- Frei, A., Bava, F., Zunder, E., Hsieh, W., Chen, Y., Nolan, G., and Gherardini, P. (2016). Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat. Methods* **13**, 269–277.
- Gaspar, I., and Ephrussi, A. (2015). Strength in numbers: quantitative single-molecule RNA detection assays. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 135–150.
- Gerdas, M.J., Sevinsky, C.J., Sood, A., Adak, S., and Bello, M.O. (2013). Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. USA* **110**, 11982–11987.
- Giesen, C., Wang, H.A., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P.J., Grolimund, D., Buhmann, J.M., Brandt, S., et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422.
- Groom, J.R., and Luster, A.D. (2011). CXCR3 in T cell function. *Exp. Cell Res.* **317**, 620–631.
- Györfi, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q., and Szallasi, Z. (2010). An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **123**, 725–731.
- Jingyi, J.L., and Biggin, M.D. (2015). Statistics requantitates the central dogma. *Science* **347**, 1066–1067.
- Johannessen, C.M., Boehm, J.S., Kim, S.Y., Thomas, S.R., Wardwell, L., Johnson, L.A., Emery, C.M., Stransky, N., Cogdill, A.P., Barretina, J., et al. (2010). COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. *Nature* **468**, 968–972.
- Kamentsky, L., Jones, T.R., Fraser, A., Bray, M.A., Logan, D.J., Madden, K.L., Ljosa, V., Rueden, C., Eliceiri, K.W., and Carpenter, A.E. (2011). Improved structure, function and compatibility for cellprofiler: modular high-throughput image analysis software. *Bioinformatics* **27**, 1179–1180.
- Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., and Nilsson, M. (2013). In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 1–6.
- Khan, I.A., MacLean, J.A., Lee, F.S., Casciotti, L., DeHaan, E., Schwartzman, J.D., and Luster, A.D. (2000). IP-10 is critical for effector T cell trafficking and host survival in *Toxoplasma gondii* infection. *Immunity* **12**, 483–494.
- Larsson, C., Grundberg, I., Söderberg, O., and Nilsson, M. (2010). In situ detection and genotyping of individual mRNA molecules. *Nat. Methods* **7**, 395–397.
- Lee, J.H., Daugherty, E.R., Scheiman, J., Kalhor, R., Amamoto, R., Peters, D.T., Turczyk, B.M., Marblestone, A.H., Yang, J.L., Ferrante, T.C., et al. (2014). Sequencing in situ. *Science* **343**, 1360–1363.
- Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.A.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197.
- Lin, J.-R., Fallahi-Sichani, M., and Sorger, P.K. (2015). Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat. Commun.* **6**, 8390.
- Liu, M., Guo, S., and Stiles, J.K. (2011). The emerging role of CXCL10 in cancer. *Oncol. Lett.* **2**, 583–589.
- Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550.
- Luster, A.D., and Ravetch, J.V. (1987). Biochemical characterization of a γ interferon-inducible cytokine (IP-10). *J. Exp. Med.* **166**, 1084–1097.
- Mackay, C.R., Terpe, H.J., Stauder, R., Marston, W.L., Stark, H., and Günther, U. (1994). Expression and modulation of CD44 variant isoforms in humans. *J. Cell Biol.* **124**, 71–82.
- Mavropoulos, A., Allo, B., He, M., Park, E., Majonis, D., and Ornatsky, O. (2017). Simultaneous detection of protein and mRNA in Jurkat and KG-1a cells by mass cytometry. *Cytometry A*. <https://doi.org/10.1002/cyto.a.23281>.
- Mulligan, A.M., Raitman, I., Feeley, L., Pinnaduwa, D., Nguyen, L.T., O'Malley, F.P., Ohashi, P.S., and Andrusis, I.L. (2013). Tumoral lymphocytic infiltration and expression of the chemokine CXCL10 in breast cancers from the Ontario Familial Breast Cancer Registry. *Clin. Cancer Res.* **19**, 336–346.
- Österreicher, C.H., Penz-Österreicher, M., Grivninkov, S.I., Guma, M., Koltsova, E.K., Datz, C., Sasik, R., Hardiman, G., Karin, M., and Brenner, D.A. (2011). Fibroblast-specific protein 1 identifies an inflammatory subpopulation of macrophages in the liver. *Proc. Natl. Acad. Sci. USA* **108**, 308–313.
- Peng, D., Kryczek, I., Nagarsheth, N., Zhao, L., Wei, S., Wang, W., Sun, Y., Zhao, E., Vatan, L., Szeliga, W., et al. (2015). Epigenetic silencing of TH1-type chemokines shapes tumour immunity and immunotherapy. *Nature* **521**, 1–16.
- Pilling, D., Fan, T., Huang, D., Kaul, B., and Gomer, R.H. (2009). Identification of markers that distinguish monocyte-derived fibrocytes from monocytes, macrophages, and fibroblasts. *PLoS One* **4**, 31–33.
- Schapiro, D., Jackson, H.W., Raghuraman, S., Fischer, J.R., Zanotelli, V.R.T., Schulz, D., Giesen, C., Catena, R., Varga, Z., and Bodenmiller, B. (2017). histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* **14**, 873–876.
- Sommer, C., Straehle, C., Koethe, U., and Hamprecht, F.A. (2011). Ilastik: interactive learning and segmentation toolkit. 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 230–233.

- Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* 14, 865–868.
- Stratton, M.R., Campbell, P.J., and Andrew, F.P. (2009). The cancer genome. *Nature* 458, 719–724.
- Takahashi, C., Au-Yeung, A., Fuh, F., Ramirez-Montagut, T., Bolen, C., Mathews, W., and O’Gorman, W.E. (2016). Mass cytometry panel optimization through the designed distribution of signal interference. *Cytometry A* 91, 39–47.
- Vassilakopoulou, M., Togun, T., Dafni, U., Cheng, H., Bordeaux, J., Neumeister, V.M., Bobos, M., Pentheroudakis, G., Skarlos, D.V., Pectasides, D., et al. (2014). In situ quantitative measurement of HER2mRNA predicts benefit from trastuzumab-containing chemotherapy in a cohort of metastatic breast cancer patients. *PLoS One* 9, 1–8.
- Visser, M., Jiwa, M., Horstman, A., Brink, A.A., Pol, R.P., Van Diest, P., Snijders, P.J., and Meijer, C.J. (2008). Intra-operative rapid diagnostic method based on CK19 mRNA expression for the detection of lymph node metastases in breast cancer. *Int. J. Cancer* 122, 2562–2567.
- Wang, F., Flanagan, J., Su, N., Wang, L.-C., Bui, S., Nielson, A., Wu, X., Vo, H.-T., Ma, X.-J., and Luo, Y. (2012). RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* 14, 22–29.
- Wang, Z., Portier, B.P., Gruver, A.M., Bui, S., Wang, H., Su, N., Vo, H.T., Ma, X.J., Luo, Y., Budd, G.T., and Tubbs, R.R. (2013). Automated quantitative RNA in situ hybridization for resolution of equivocal and heterogeneous ERBB2 (HER2) status in invasive breast carcinoma. *J. Mol. Diagn.* 15, 210–219.
- Weatherly, K., Bettonville, M., Torres, D., Kohler, A., Goriely, S., and Braun, M.Y. (2015). Functional profile of S100A4-deficient T cells. *Immun. Inflamm. Dis.* 3, 431–444.
- Wightman, S.C., Uppal, A., Pitroda, S.P., Ganai, S., Burnette, B., Stack, M., Oshima, G., Khan, S., Huang, X., Posner, M.C., et al. (2015). Oncogenic CXCL10 signalling drives metastasis development and poor clinical outcome. *Br. J. Cancer* 113, 1–9.
- Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., Chambers, M.C., Zimmerman, L.J., Shaddox, K.F., Kim, S., et al. (2014). Proteogenomic characterization of human colon and rectal cancer. *Nature* 513, 382–387.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Histone H3	Cell Signaling Technologies (CST)	Cat#: 4499; RRID: AB_10544537
Cytokeratin 5	Abcam	Cat#: ab52635; RRID: AB_869890
Cytokeratin 8/18	CST	Cat#: 4546; RRID: AB_2134843
Cytokeratin 14	ThermoFisher	Cat#: PA5-16722; RRID: AB_10980222
Cytokeratin 19	Dev. Studies Hybridoma Bank	Troma-III; RRID: AB_2133570
SMA	Abcam	Cat#: ab76549; RRID: AB_2223019
Vimentin	CST	Cat#: 5741; RRID: AB_10695459
CD68	E-Bioscience	Cat#: 14-0688-82; RRID: AB_11151139
CD3	CST	Cat# 85061
CD44	R&D Systems	Cat#: AF3660; RRID: AB_10971655
HER2	CST	Cat#: 4290; RRID: AB_10828932
FSP1	Millipore	Cat#: 07-2274; RRID: AB_10807552
E-/N-cadherin	Becton Dickinson	Cat#: 610182; RRID: AB_397581
Pan-Keratin AE1	Millipore	Cat#: MAB1612; RRID: AB_2132794
Pan-Keratin AE3	Millipore	Cat#: MAB1611; RRID: AB_2134409
Pan-Keratin	Dako	Cat#: Z0622; RRID: AB_2650434
pS6	CST	Cat#: 4858; RRID: AB_916156
Ki-67	CST	Cat#: 9449
Bacterial and Virus Strains		
Biological Samples		
Breast cancer tissue	Univeristy Hospital Zurich	N/A
Breast cancer tissue micro array	Yale School of Pathology	N/A
Chemicals, Peptides, and Recombinant Proteins		
Ir Intercalator	Fluidigm	Cat# 201192A
Critical Commercial Assays		
MaxPar labeling kit	Fluidigm	Cat# 201300
RNAscope fluoresecent multiplex kit	ACD	N/A
Spin columns for oligonucleotide purification	Millipore	Cat# MRCF0R030
Chamber slides	Ibidi	Cat# 81201
Deposited Data		
Images from breast cancer TMA		https://doi.org/10.17632/m4b97v7myb.1
Images from IMC and IF comparison		https://doi.org/10.17632/m4b97v7myb.1
Images from buffer comparison		https://doi.org/10.17632/m4b97v7myb.1
Images from HeLa cell pellets		https://doi.org/10.17632/m4b97v7myb.1
Images from control gene expression		https://doi.org/10.17632/m4b97v7myb.1
Experimental Models: Cell Lines		
A431 cell line	University of Zurich	Pelkman's lab
HeLa cell pellets	ACD	Cat# 310045
Oligonucleotides		
CK19 probes channel 1	ACD	Cat# 310221
CXCL10 probes channel 2	ACD	Cat# 311851-C2
ERBB2 probes channel 3	ACD	Cat# 310081-C3
POLR2A probes channel 1	ACD	Cat# 310451

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PPIB probes channel 2	ACD	Cat# 313901-C2
UBC probes channel 3	ACD	Cat# 310041-C3
GAPDH probes channel 1	ACD	Cat# 310321
ACTB probes channel 2	ACD	Cat# 310141-C2
dapB probes channel 1/2/3	ACD	Cat# 310043-C1/2/3
TBP probes channel 2	ACD	Cat# 314291-C2
Recombinant DNA		
CK19 overexpression vector	GenScript	Clone ID: OHu19117
CXCL10 overexpression plasmid	Dharmacon	Clone ID: 6354
ERBB2 overexpression vector	Addgene	Plasmid #23888
pDEST pcDNA5 FRT TO-eGFP	Couzens et al., 2013; University of Toronto	N/A
Software and Algorithms		
CellProfiler	Kamentsky et al., 2011	www.cellprofiler.org
Ilastik 1.1.19	Sommer et al., 2011	www.ilastik.org
HistoCAT	Schapiro et al., 2017	www.bodenmillerlab.org
R 3.4.1	R Development Core Team, 2015	https://www.R-project.org
Matlab 2014b	MathWorks	www.mathworks.com
Python 3.4.3	Python Software Foundation.	www.python.org
PhenoGraph	Levine et al., 2015	R implementation: https://github.com/JinmiaoChenLab/Rphenograph

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Bernd Bodenmiller (bernd.bodenmiller@imls.uzh.ch).

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Biological Material**

Human FFPE HeLa cell pellet sections with 5 μ m thickness were obtained from ACD. Human breast cancer tissue sections were obtained from our collaborators at the University Hospital Zurich under ethics approval KEK-ZH-no.: 2012-0553. The HER2 TMA (FFPE) was previously described (Carvajal-Hausdorf et al., 2015) and was kindly provided by David Rimm from the Yale School of Pathology. Overexpression experiments were performed in A431 breast cancer cell lines. Cells were cultured in DMEM (D5671, SIGMA), supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin, and 100 μ g/ml streptomycin. A431 cells were only used for overexpression of mRNA target genes and no authentication of the cell line was performed in this study.

METHOD DETAILS**Experimental Design**

The number of independent replicates is given in the respective figure legends. No sample size estimations and no blinding were performed. The cores from the HER2 TMA were acquired in a random order to exclude bias due to long measurement times. Further information on statistical tests is provided in the respective [STAR Methods](#) section.

Antibody Conjugation

Carrier-free antibodies were conjugated to metal tags using the MaxPar® labeling kit (Fluidigm) following the manufacturer's instructions. The yield of each conjugation was determined photometrically, and antibodies were stored at concentrations ranging from 100-400 μ g/ml in stabilizing solution (Candor Biosciences) at 4°C.

Oligonucleotide Conjugation

Oligonucleotides were conjugated to metal tags as previously described (Frei et al., 2016). Briefly, thiol-modified C6 S-S oligonucleotides were obtained from ACD. Oligonucleotides (7 mmol) in ddH₂O were reduced with TCEP (final concentration 50 mM) for 30 min at room temperature followed by an ethanol precipitation. The pellets were resuspended in 50 μ l C-Buffer (MaxPar labeling Kit). Simultaneously the MaxPar polymers were loaded with metals and purified in C-Buffer according to the manufacturer's protocol. For the oligonucleotide polymer conjugation 2 nmol of oligonucleotides were conjugated with 100 μ g polymer in C-Buffer for 2 h at room temperature. After 2 h, 5 nM TCEP was added to the reaction to reduce oligonucleotide dimers. The reaction was purified with 30-kDa filter columns (MRCF0R030, Millipore) and washed twice with ddH₂O, and the conjugated oligonucleotides were resuspended in 100 μ l ddH₂O. Conjugated and unconjugated oligonucleotides were run on a 4% agarose gel and visualized using GelRed (Biotium). The concentration of the oligonucleotides was determined photometrically, and a 1 μ M dilution in ddH₂O was made and stored at 4°C.

Sample Pre-treatment

Prior to RNA staining with RNAscope® Fluorescent Multiplex Kit, samples were pre-treated according to manufacturer's protocol for FFPE samples. For the comparison with the HIER protocol consecutive sections were cut at four micrometer thickness from a FFPE tissue block, and the sections were subsequently baked on Superfrost Plus slides at 37°C for 2 h. Sections were deparaffinized in xylol (2x10 min). After a 10-min wash in xylol/100% ethanol (1:1) sections were rehydrated in a graded alcohol series (100%, 90%, 80%, 70%, 50%) for 10 min each step, followed by a wash in TBS for 10 min. Antigen retrieval was performed using Tris-EDTA buffer (pH 9.2) for 20 min at 95°C. After cooling, blocking buffer (3% normal horse serum, 0.5% Triton-X-100 in TBS) was applied for at least 1 h at room temperature. Sections were incubated with antibodies overnight at 4°C. After brief washing samples were incubated with MaxPar Intercalator-Ir (1:1000; Fluidigm, Cat: 201192B) in PBS for 5 min, rinsed (3x5 min) with TBS, dried under airflow, and stored at room temperature until measurements.

IMC and FISH Comparison

To compare IMC and FISH measurements we used HeLa cell pellets. The RNA staining was performed with the RNAscope Fluorescent Multiplex Kit following the manufacturer's protocols. During the final staining step of the RNAscope protocol for the hybridization of the detection oligonucleotides a mix of 10 nM metal-labeled and fluorophore-labeled oligonucleotides was applied. The sections were then briefly washed in TBS and then stained with antibodies in TBS-Tween overnight at 4°C. The next day the slides were stained with DAPI for 5 min followed by a washing step in PBS and then stained with a 1:1000 dilution of 500 μ M MaxPar Intercalator-Ir (Fluidigm) in PBS for 5 min. The samples were then washed twice with PBS and imaged for fluorescence staining on a Zeiss Axio Scan.Z1 slidescanner (ORCA-Flash 4.0 camera; objective, Plan-Apochromat 40x/0.95; filter sets, DAPI, Alexa Fluor 488, and Alexa Fluor 548). Images were acquired using Zen blue software, and data was exported as 16 bit tiff images. After fluorescence imaging the coverslips were carefully removed and the slides dried under airflow for IMC measurements. On the IMC an area that had previously been imaged with the fluorescent microscope was selected for ablation and metal detection.

RNA and Protein Staining

RNA staining was performed with an RNAscope Fluorescent Multiplex Kit following the manufacturer's protocols. Metal-labeled oligonucleotides were used at a final concentration of 20 nM. For simultaneous antibody staining, the sample slides were briefly washed with TBS after the last wash in the RNAscope protocol. Subsequently slides were stained with antibodies (diluted in TBS-Tween) overnight at 4°C in a wet chamber. The next day slides were washed for 5 min in TBS and then stained for 5 min in a 1:1000 dilution of 500 μ M MaxPar Intercalator-Ir (Fluidigm) in PBS. Slides were washed for 5 min in PBS, dried under airflow, and stored at room temperature until measurements.

Overexpression of *CK19* and *CXCL10* and *ERBB2*

CK19 expression-ready vector was obtained from GenScript (Clone ID: OHu19117). *ERBB2* (Plasmid #23888, Addgene) (Johannesen et al., 2010) and *CXCL10* (Clone ID: 6354, Dharmacon) were cloned into a pDEST pcDNA5 FRT TO-eGFP vector (Couzens et al., 2013) via Gateway Cloning for transfection. A431 cells (a kind gift from the Pelkmans lab, University of Zurich) were seeded at the density of 50,000 per well in a 12-well chambered slide (Ibidi) 24 h before transfection and were transfected with 0.3 μ g plasmid and 0.6 μ l of jetPRIME (PolyPlus) per well with the standard protocol provided by the manufacturer. After 16–20 h, cells were briefly washed with PBS and then fixed for 15 min with 10% formaldehyde. Cells were washed with PBS and then permeabilized with 0.1% Triton in PBS for 5 min. Cells were then treated with a 1:15 dilution of Protease III (ACD) at room temperature and washed once with PBS and twice with ddH₂O. Probe hybridization was performed using the ACD multiplexed fluorescence kit following the manufacturer's protocol.

Data Acquisition

For IMC, images were acquired with a beta prototype unit of the laser module from Fluidigm coupled to a Helios mass cytometer. As the machine is under development various software developments and changes accompanied the experiments. Early versions required manual tuning of the argon and helium gas flows before every acquisition in order to achieve optimal carriage of ablated materials to the plasma of the mass cytometer. Therefore, a metal-coated tuning slide (Fluidigm) was used to optimize peak intensity

and resolution as a function of helium and argon gas flow. Later software versions included automated tuning procedures performed on a daily basis before acquisitions. The TMA data were acquired in a single batch over one week. All images were acquired at 200 Hz, and data are stored as MCD files as well as txt files.

QUANTIFICATION AND STATISTICAL ANALYSIS

Image Generation

The txt files output from the IMC instrument were converted to tiff images using custom written code in Python (<https://github.com/BodenmillerGroup/imctools>). False color images for visualization were prepared from the tiff images in FIJI open source software.

Pixel Classification in Ilastik for Segmentation

For single-cell analysis, tiff stacks of images from histone H3, DNA1, DNA2, panCytokeratin, E-/N-cadherin, and HER2 were created. These stacks were used for pixel classification in Ilastik open source software (Sommer et al., 2011). Using Ilastik we defined pixels as belonging to the nuclear, cytoplasmic, or background compartments and exported these class probabilities as RGB tiff images. For the TMA analysis randomly selected 125 x 125 pixel sections of each of the channels analyzed were generated from the tiff stacks and then scaled up to generate 250 x 250 pixel images, which are easier to classify. The pixel classification was trained on the random small image stacks of all samples and was subsequently applied to the original large image stacks in batch mode to generate probabilities for nuclear, cytoplasmic, and background signals. These probability images were then exported as RGB tiff files.

Single Cell Segmentation

Ilastik RGB probability images were subsequently imported into CellProfiler (Kamentsky et al., 2011) along with the original images (all channels) and used for single-cell segmentation, mask generation, and marker quantification. From the probabilities from Ilastik we defined nuclei as primary objects. To define cells we expanded the nuclei toward the border between cytoplasm and background as defined from the Ilastik cytoplasm and background channel and then scaled them down to the original resolution. A mask for the identification of single cells was generated in CellProfiler and was saved as a tiff image. The abundance of either RNA or antibody signals for every cell was also quantified in CellProfiler and exported as a table for further use. Image segmentation from the over-expression experiments was performed using the “IdentifyObjectsManually” CellProfiler module. Single cell quantifications for roughly 50-100 cells per condition were exported for plotting in R.

Data Analysis

The generated single-cell data were then further analyzed using R and histoCAT (Schapiro et al., 2017). Log10 transformation was applied as indicated in the figures. However, because in the RNA channels we observed cells with a mean of 0 counts we set all cells with mean counts lower than 0.005 to a minimum value of 0.005. Quantile normalization (99th percentile) was applied for PhenoGraph clustering (parameter k = 50) as proposed in the original publication (Levine et al., 2015). All correlation based analyses were performed using Spearman correlation, which is invariant to data transformation or Pearson correlation on untransformed data for the buffer comparison.

Statistical Tests and Data Exclusion

The TMA consisted of 88 samples, but several samples were lost during the TMA processing. Samples with fewer than 200 total cells were excluded from the analysis which resulted in a final dataset of 70 samples. Of the 70 samples, three samples had no epithelial cells and were also excluded. For the statistical analysis of the *HER2* mRNA and protein expression in epithelial cells compared to epithelial cells from control tissues we used a one tailed t-test. Because patient samples were only analyzed once, we assumed equal variance as observed for the six different control tissue samples on the TMA. No significances were calculated for samples with less than 50 epithelial cells, and correlations for RNA and protein were only calculated for samples with at least 50 epithelial cells. For RNA integrity estimates samples with less than 50 epithelial cells were discarded. This resulted in 10 samples with 95% or less of the epithelial cells having less than -0.5 transformed mean counts (samples C10, D03, E05, E10, F04, F06, F10, H10, I01, J10). Of these, samples D03, C10 and J10 did not express either CK19 or HER2 protein. This resulted in 7 samples corresponding to 11% of all samples with more than 50 epithelial cells that showed protein expression but not RNA expression and may thus have suffered from RNA degradation. The statistical significance of the association between observing T cells (<10%) in the absence of *CXCL10*^{high} cells (<1%) was determined using a Barnard’s test on the contingency table.

Neighborhood Analysis Using histoCAT

For the standard neighborhood permutation test our recently published software histoCAT was adapted to allow studies of the neighborhood of manually gated populations as well as specifically test for patches. This modified version of histoCAT (patch detection) is available on GitHub. For the *CXCL10*^{high} neighborhood, cells that expressed high levels of *CXCL10* were gated manually within histoCAT. The algorithm was adapted to calculate the occurrence of cells with two or more neighbors of the same cell phenotype, which we define as a “patchy” motif. Comparing the frequency of this patchy motif in an image versus the empirical distribution of

999 random permutations of the cell class labels in the same image, while maintaining the overall connectivity structure of the image, allowed calculation of p-values. These indicate how likely it would be to observe as many or more patchy motifs in an image by chance.

RNA FISH and IMC Data Analysis

The tiff images from the slide scanner and from the IMC were cropped to be roughly the same area and then overlaid in FIJI using the turboreg plugin (bilinear, manual setting of landmarks). Landmarks were exported and IMC images scaled up using software custom written in Matlab to reach the same resolution as microscopy images. For the upscaling of the IMC images a similarity transform was used and the total intensity per image was kept constant resulting in a factor by which the absolute IMC signal was increased through upscaling. Pixel classification and segmentation were performed as described above to generate the single-cell mask. For mRNA spot counting in the fluorescence images a previously published spot counting algorithm (Battich et al., 2013) was implemented in our published software histoCAT (Schapiro et al., 2017). This newest release of histoCAT now screens for “RNA” in the names of the images to be loaded and if true will automatically start the spot counting module. Parameters for spot detection match those from the previous publication (Battich et al., 2013) and were set as follows: Object size = 5; Intensity Quanta Per Image = 0.02 0.99; Threshold of Spot Detection = 0.013; Steps of Deblending = 30. histoCAT allows the user to inspect the spot detection and adapt the parameters to rerun the spot detection before continuing. Once all the images and the mask are loaded histoCAT also automatically creates the “Spots” channel for the image that contained “RNA” in the filename. Additionally, histoCAT calculates the absolute ion counts per cell for all IMC channels. This allows direct comparison of either mean fluorescence intensity per cell versus mean ion counts per cell or it allows comparison of absolute RNA spots per cell to absolute ion counts per cell. Of note, the ion counts in histoCAT are multiplied by the factor of image upscaling. Therefore, to obtain the correct ion counts per single-cell the data was exported from histoCAT, IMC data divided by the upscaling factor and plotted in R.

DATA AND SOFTWARE AVAILABILITY

The accession number for the data reported in this paper is MendeleyData: <https://doi.org/doi:10.17632/m4b97v7myb.1>. The spot detection and patch detection implemented in histoCAT is available in the newest release of histoCAT (<https://github.com/BodenmillerGroup/histoCAT>). All custom code from Matlab, Python, and R is available upon request.